SMAI: Assignment2 (Emotion Prediction)

Upendra Kumar
IIIT Hyderabad

Abstract

In this report, the challenging problem of visual emotion analysis is attempted. Multiple feature sets obtained from PCA, LDA, KernelPCA, KernelLDA, VGG and Resnet are extracted and used for identifying emotion from images.

1 Introduction

Nowadays, people share a lot of content on social media in the form of images. Analyzing image content from social media websites and photo-sharing websites like Flickr, Twitter, Tumblr, etc., can give us valuable insights into the general sentiment of people about say a legislation, popularity of an actor or movies. Therfore, visual emotion analysis is an important problem.

Possible application of visual emotion analysis can be in content filtering systems which can filter aggressive emotions or in a social media analysis platform where it is necessary to get overall sentiment of people about a particular aspect.

2 System Components

Features used : PCA, LDA, KernelPCA, KernelLDA, VGG and Resnet
Dataset used : IMFDB and Yale Image dataset

3 Experimental Setup

In this section, we describe our setup in detail. Experimentation was done with set of three classifiers : Logistic Regression, Multi Layer Perceptron and SVM with different combinations of features.

4 Quantitative Results

	feature	dimension	classification error	accuracy	precision
0	Eigen Face	100	0.45	0.55	0.531704
1	Kernel Fisher Face	6	0.09	0.91	0.896009
2	Fisher Face	6	0.09	0.91	0.896009
3	Vgg Features	4096	0.44	0.56	0.574685
4	Kernel Face	100	0.62	0.38	0.144400
5	Resnet Features	2048	0.52	0.48	0.477688

Table 1: Dataset imfdb Classifier logistic

	feature	dimension	classification error	accuracy	precision
0	Eigen Face	100	0.50	0.50	0.504528
1	Kernel Fisher Face	6	0.11	0.89	0.897158
2	Fisher Face	6	0.11	0.89	0.897158
3	Vgg Features	4096	0.55	0.45	0.481170
4	Kernel Face	100	0.40	0.60	0.626604
5	Resnet Features	2048	0.54	0.46	0.478647

Table 2: Dataset imfdb Classifier mlp

	feature	dimension	classification error	accuracy	precision
0	Eigen Face	100	0.54	0.46	0.390093
1	Kernel Fisher Face	6	0.08	0.92	0.923208
2	Fisher Face	6	0.08	0.92	0.923208
3	Vgg Features	4096	0.40	0.60	0.602150
4	Kernel Face	100	0.53	0.47	0.402464
5	Resnet Features	2048	0.37	0.63	0.677047

Table 3: Dataset imfdb Classifier svm

	feature	dimension	classification error	accuracy	precision
0	Eigen Face	100	0.738095	0.261905	0.361111
1	Kernel Fisher Face	10	0.142857	0.857143	0.888889
2	Fisher Face	10	0.142857	0.857143	0.888889
3	Vgg Features	4096	0.952381	0.047619	0.130952
4	Kernel Face	100	0.904762	0.095238	0.029630
5	Resnet Features	2048	1.000000	0.000000	0.000000

Table 4: Dataset yale Classifier logistic

	feature	dimension	classification error	accuracy	precision
0	Eigen Face	100	0.857143	0.142857	0.180272
1	Kernel Fisher Face	10	0.190476	0.809524	0.858333
2	Fisher Face	10	0.190476	0.809524	0.858333
3	Vgg Features	4096	0.976190	0.023810	0.059524
4	Kernel Face	100	0.738095	0.261905	0.341043
5	Resnet Features	2048	0.976190	0.023810	0.007937

Table 5: Dataset yale Classifier mlp

	feature	dimension	classification error	accuracy	precision
0	Eigen Face	100	0.857143	0.142857	0.194444
1	Kernel Fisher Face	10	0.142857	0.857143	0.869048
2	Fisher Face	10	0.142857	0.857143	0.869048
3	Vgg Features	4096	0.976190	0.023810	0.007937
4	Kernel Face	100	0.880952	0.119048	0.194678
5	Resnet Features	2048	1.000000	0.000000	0.000000

Table 6: Dataset yale Classifier svm

5 Examples

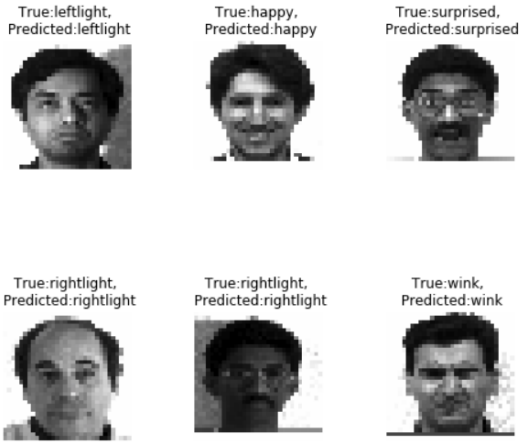

Figure 1: Examples of Emotion Prediction from Yale Dataset

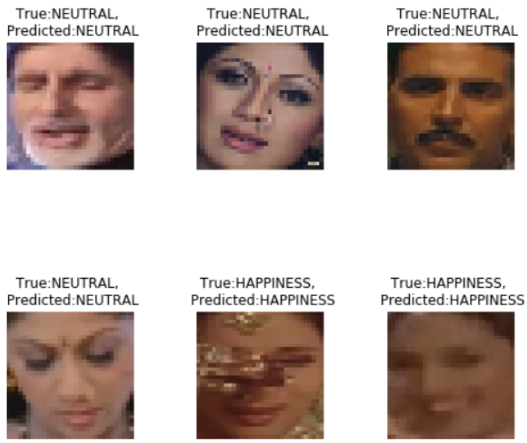

Figure 2: Examples of Emotion Prediction from IMFDB Dataset

6 Conclusion

In this report, only basic machine learning techniques of feature extraction and training models were used. Even if the results are encouraging, but still a lot of issues are left unaddressed. For example, performance of model trained on IMFDB dataset on YALE Dataset and vice-versa. This kind of analysis can reveal how much model can adpat ot domain. Of course in our case, since data for training is limited, generalization capability of model will be very less. That is model is expected to perform poorly on images obtained from other domain. These problems can be mitigated by use of larger dataset and better models for obtaining features (like end-to-end deep learning models for predicting emotion and specialized models for detecting facial features).

As an interesting future work, one can consider the presence of a person in the image as well as the facial expression. Facial expressions can have high correlation with emotion. Even if the overall mood of the image is dystopian, a cheerful face can mitigate negative emotion.

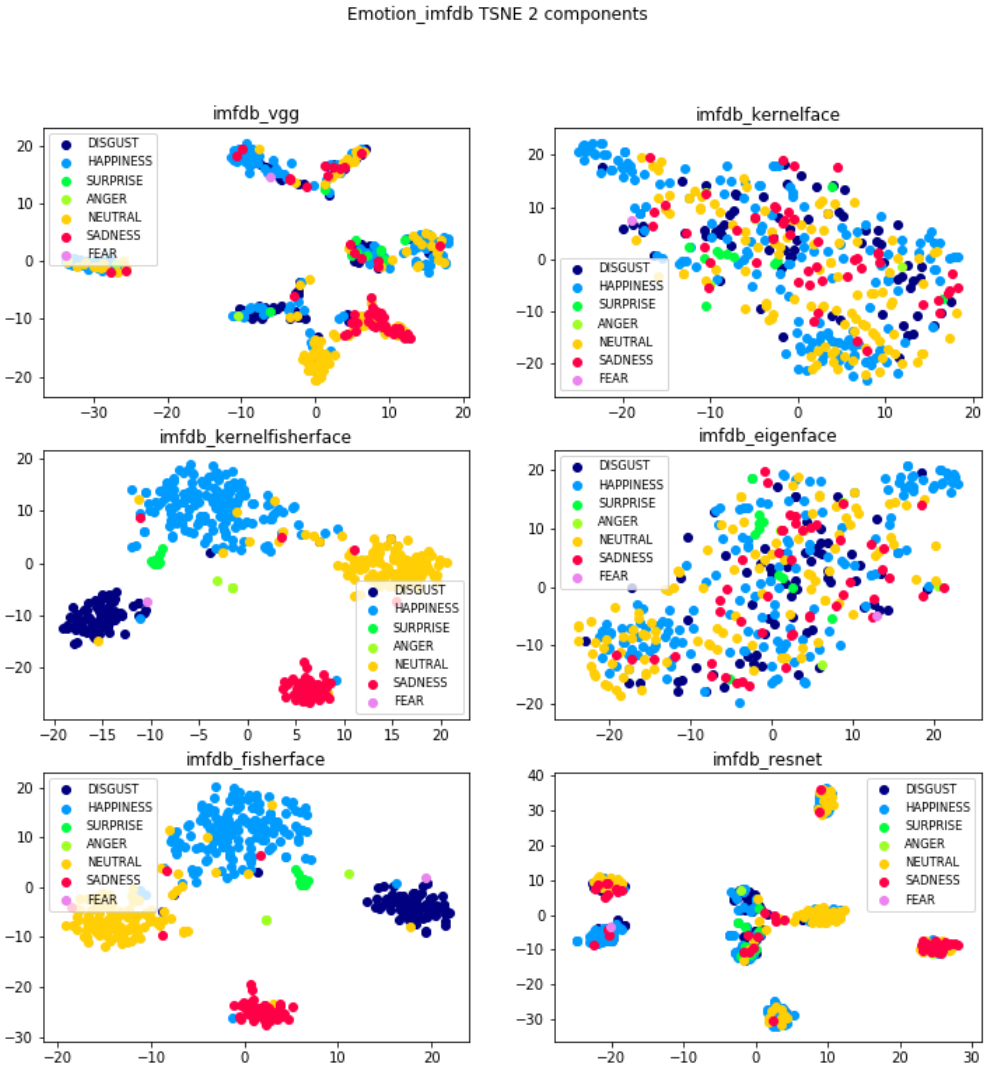

Figure 3: TSNE Visualization of features for emotion detection from IMFDB Dataset

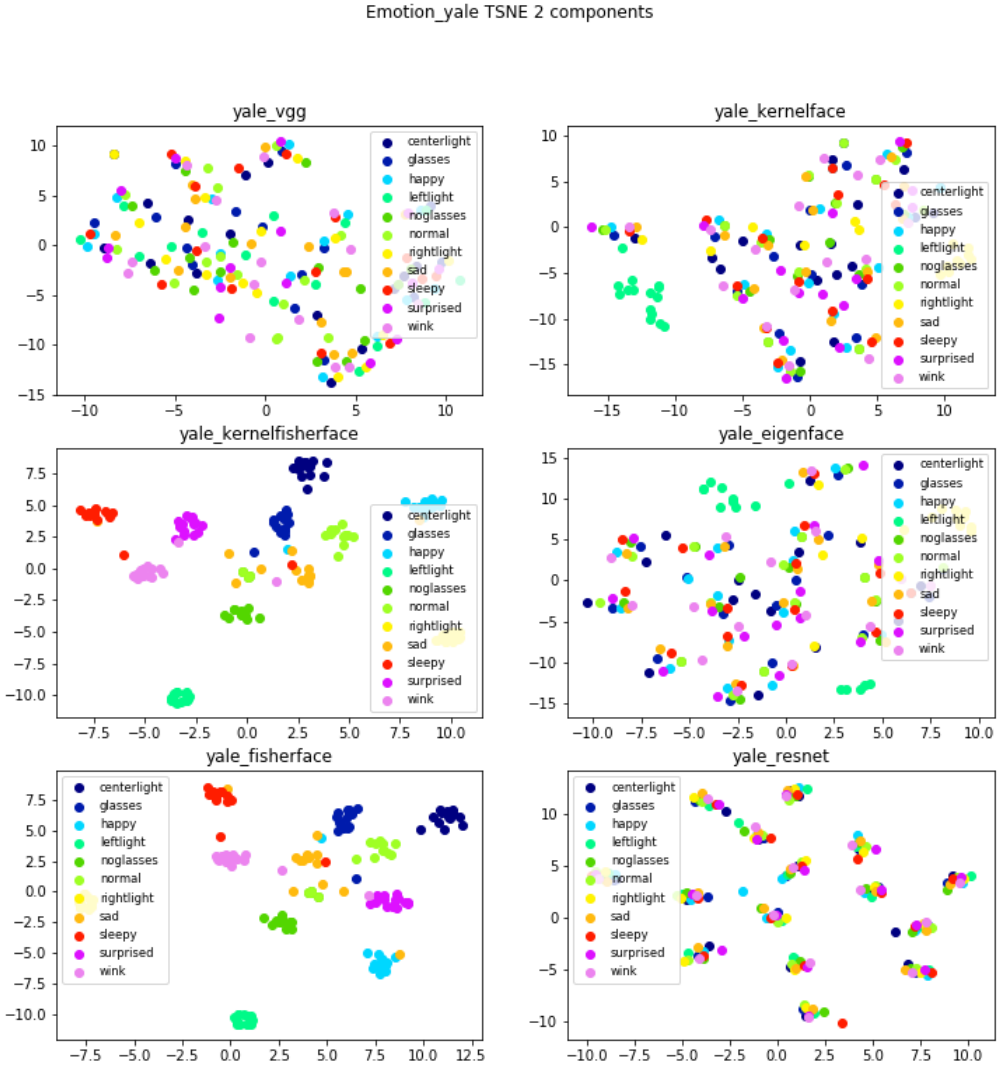

Figure 4: TSNE Visualization of features for emotion detection from Yale Dataset